

# **ПРОБЛЕМЫ ПРЕДСТАВЛЕНИЯ, ПРИОБРЕТЕНИЯ И ИСПОЛЬЗОВАНИЯ ЗНАНИЙ В СВЕТЕ ОБРАБОТКИ ЕСТЕСТВЕННОГО ЯЗЫКА**

**В.Н. Поляков**

**МГЛУ, Кафедра экспериментальной и прикладной  
лингвистики**

[tan@ibpm.serpukhov.su](mailto:tan@ibpm.serpukhov.su), [vladimir\\_polyakov@yahoo.com](mailto:vladimir_polyakov@yahoo.com)

<http://www.geocities.com/SiliconValley/Campus/7926/Polyakov/Polyakov.htm>

**Аннотация.** В работе исследованы пути дальнейшего развития исследований в области представления, приобретения и использования знаний. В качестве отправной точки принята проблема обработки естественного языка. Рассмотрены достоинства и недостатки моделей представления знаний и смысла ЕЯ-форм, занимающих лидирующие позиции в современных интеллектуальных информационных системах. В качестве пути преодоления указанных недостатков предлагается семиотическая модель понимания текста. Понимание текста рассматривается как сложный процесс, связанный с интерпретацией написанного (сказанного) текста с учетом сосуществования целого комплекса семиотических подсистем. К таким подсистемам относятся: непосредственно фрагмент текста; знания агента о языке; ситуация, в которую погружено повествование; картина мира агента, интерпретирующего текст; цели агента и связанный с этими целями репертуар моделей поведения (или моделей решения задачи, обеспечивающих достижение цели). В работе предложена концепция частичного понимания текста. Для приобретения знаний в рамках этой концепции описан цикл обучения, позволяющий объединить процесс автоматического формирования картины мира и знаний о языке по корпусу текстов. Для целей использования знаний предложен двухуровневый механизм вывода, который объединяет элементы систем представления знаний и систем принятия решений (экспертных систем) и удовлетворяет следующим условиям: подчиненность деятельности определенной цели; использование опыта; использование знаний. Работу можно рассматривать как программу научных исследований, которую, по мнению автора, предстоит реализовать научному сообществу в ближайшее десятилетие.

## 1. ВВЕДЕНИЕ

Искусственный интеллект (ИИ) [19] прочно занял свое место в области информационных технологий как средство интеллектуализации информационных систем. В последнее десятилетие наметилась тенденция интеграции моделей ИИ в рамках перспективных технологических направлений. При этом наблюдается синергетический эффект, позволяющий получить новое качество. Примерами такого синтеза моделей являются:

- технология "мягких вычислений", объединяющая нейросети, нечетко-значные и модальные логики;
- технология гибридных экспертных систем [4,16], объединяющая традиционные модели программирования и продукционные модели представления знаний.

На наш взгляд перспектива построения интеллектуальных информационных систем лежит на пути комплексного использования:

- моделей представления знаний [19, т.2];
- моделей и методов обработки естественного языка (ОЕЯ)[32];
- методов логического вывода [33];
- моделей принятия решений на основе экспертных систем [19, т.1].

Однако существует ряд нерешенных фундаментальных проблем, ограничивающих широкое применение технологий представления, приобретения и использования знаний в образовательной сфере. В первую очередь это:

- отсутствие технологии создания баз знаний, использующих не только экспертные знания, но и знания о мире;
- отсутствие технологии приобретения знаний из текстов;
- отсутствие приемлемых моделей, сочетающих способы принятия решений в экспертных системах и методы логического вывода на базе когнитивных моделей предметной области.

Знания как основа интеллектуальной деятельности человека являются объектом исследования ряда дисциплин: психологии, педагогики, математики, лингвистики, когнитивной науки, медицины, биологии, гносеологии.

Искусственный интеллект изучает знания с точки зрения моделирования интеллектуальной деятельности средствами ЭВМ [14]. К настоящему времени в рамках ИИ накоплено разнообразное количество сведений, позволяющих судить о феноменологии знаний как таковых и особенностях основных процессов сопровождающих знания: представлении знаний (т.е. способе их описании в компьютере), приобретении знаний и их использовании.

Можно выделить три основных подхода к исследованию проблемы знаний. На основе подхода, сформированного в когнитивной психологии [18], исследователи занимаются фундаментальными проблемами феноменологии знаний. Лингвисты (лексикографы и когнитивные) [1,6,29], занимаются проблемами отражения знаний о мире в языке. В рамках искусственного интеллекта занимаются созданием систем, основанных на знаниях [11, т.2; 33].

В когнитивной психологии существует точка зрения о наличии двух способов представления знаний, которыми владеет человек: пропозициональная (т.е. символическая) форма и образная (т.е. перцептивная) форма.

В разделе ИИ "представление знаний" занимаются преимущественно моделями и методами обработки пропозициональной формы знаний, что в свою очередь позволяет широко использовать формализмы, основанные на математической логике и других достижениях дискретной математики. При этом логическое умозаключение выступает в качестве универсальной основы для создания методов использования пропозициональных знаний.

Образной формой знаний в самой общей постановке вопроса занимается теория распознавания образов.

Существует распространенная точка зрения [17] о разделении знаний на два типа: процедурные и декларативные. При этом декларативные знания призваны отвечать на вопрос "что", а процедурные "как". Однако такое деление весьма условно, так как любая процедура включает в себя параметры, которые уже относятся к декларативной части. В то же время процедура может быть представлена в виде описания, а не действия, и тогда ее статус размывается.

Строгих критериев отличия данных и знаний не введено. Автор статьи считает, что знания - суть сильно разветвленные данные. В [11] введен критерий разграничения знаний от просто информации (данных) на основе их деятельностного освоения, т.е. в сущности, информация приравнивается к декларативной компоненте знаний, что подтверждает точку зрения автора.

Различают также структуры для представления знаний и непосредственно знания. Первые принято называть моделями (также: формализмы, парадигмы) представления знаний, вторые - моделями знаний (также: когнитивными моделями, картиной мирой, наивными логиками, псевдофизическими логиками, логиками здравого смысла) (см. например [5]).

Наиболее распространенными моделями представления знаний являются фреймы, семантические сети и графы. В последнее время, появилось новое направление - онтологии, в рамках которого осуществляется попытка интеграции различных направлений представления знаний.

Минимальными структурами знаний принято считать объекты и отношения между объектами. Часто выделяют атрибуты как разновидность подчиненных объектов, свойства и состояния как разновидность отношений. С точки зрения уровня общности выделяют такой вид знаний как мета-знания, т.е. знания о знаниях, абстрактные знания (понятия, общие свойства и отношения, множества объектов, сценарии, типовые ситуации), и конкретные знания (индивиды, значения отношений, ситуации - прецеденты).

В когнитивной психологии рассматривается два подхода к классификации объектов: классификация, основанная на инвариантном наборе свойств и диапазоне их значений, и классификация, основанная на сходстве с прототипом, т.е. типичным представителем класса.

Существует также распространенная классификация отношений<sup>1</sup> по видам когнитивных моделей (=моделей знаний) [27]:

- отношения элемент-множество;
- отношения часть-целое;
- отношения объект-свойство;
- планы, действия и изменения;
- причинно-следственные отношения;
- темпоральные (т.е. временные) отношения;
- пространственные отношения.

По нашему мнению, кроме перечисленных, необходимо выделять еще, по крайней мере, три группы отношений:

- информационные, ментальные и коммуникативные отношения;
- социальные отношения;
- отношения человека и окружающей среды.

В последнее время в области представления знаний предприняты попытки формализации языка описания знаний. Представлены два конкурирующих проекта такого стандарта:

KIF (Knowledge Interchange Format) [26] и

CGIF (Conceptual Graphs Interchange Format) [23],

которые создают базу для унификации систем, основанных на знаниях.

---

<sup>1</sup> Под отношениями здесь понимается широкий класс семантических отношений, в том числе унарные отношения, т.е. состояния объектов.



Необходимо подчеркнуть особую роль естественного языка как универсальной формы описания знаний, которая послужила метафорической основой для многих формальных систем. Так все известные формы логических исчислений (логика высказываний, логика предикатов, нечеткозначные логики, модальные логики) были "подсмотрены" математиками в естественном языке (см.[22]).

Кроме того, в языке мы находим грамматические формы, которые служат отражением современных форм представления знаний - фреймов и семантических сетей. С другой стороны, тексты на естественном языке (книги, учебники, справочники, лекции, статьи и т.д.) служат хранилищем и источником знаний. Известно также, что текст на естественном языке - это вербализованная форма умозаключения [7].

Таким образом, знание и язык - это неразрывные сущности, и проблемы представления, приобретения и использования знаний необходимо рассматривать через призму обработки естественного языка и наоборот.

В настоящей работе приведены соображения по вопросам дальнейшего развития исследований в области представления, приобретения и использования знаний. В качестве отправной точки будущего развития данной отрасли науки видится проблема обработки естественного языка. Работу можно рассматривать как программу научных исследований, которую, по мнению автора, предстоит реализовать научному сообществу в ближайшее десятилетие.

## **2. ПРЕДСТАВЛЕНИЕ ЗНАНИЙ И СМЫСЛА ЕСТЕСТВЕННО-ЯЗЫКОВОГО ТЕКСТА**

Рассмотрим кратко достоинства и недостатки моделей представления знаний и смысла ЕЯ-форм, занимающих лидирующие позиции в современных научных изданиях по указанной тематике.

Семантические сети - это пропозициональная форма знаний, соединенная с графовым представлением [33]. Большинство формализмов семантических сетей [27] разрабатывались для целей обработки естественного языка. Однако, исследования последних лет [13] показали, что семантические сети имеют ряд существенных недостатков, затрудняющих представление смысла ЕЯ-предложения (высказывания). Главный недостаток - это отсутствие коммуникативной направленности семантического представления, что не позволяет передавать актуальное членение, фокус эмпатии, коммуникативную функцию высказывания. Кроме того,

представление в семантической сети инвариантно к контексту. В семантических сетях отсутствуют средства представления аналитических форм отношений, что не позволяет строить семантические эквиваленты толкований. Последнее замечание не относится к теории концептуальных зависимостей (ТКЗ) [31], в которой концептуализации представляют собой некоторые разновидности толкований. Кроме того, в рамках ТКЗ было предложено несколько решений, представляющих интерес с точки зрения интерпретации ЕЯ-форм. Это:

- мета-структуры (SCRIPTS, т.е. сценарии, memory organisation packets (MOP) и др.), ставшие прообразами CASE-структур;
- психологически-ожидаемое умозаключение;
- семантически-ориентированный формализм грамматического разбора, основанный на методе гипотез и ожиданий.

Однако при всех своих плюсах ТКЗ имеет ряд недостатков, которые не позволили ей занять лидирующее положение в соревновании формализмов семантических сетей между собой. К ним относится следующее:

- ТКЗ искусственно дистанцирована от теории графов и математической логики, хотя оперирует сходными понятиями;
- ТКЗ оторвана от лексикографических исследований, хотя близка им по сути.

Теория фреймов (ТФ) [9] - это соединение пропозициональной и объектно-ориентированной формы представления знаний. Ограничением применения ТФ является то, что она охватывает достаточно узкий класс отношений - свойств. Кроме того, современные исследования в когнитивной психологии [18] показали, что для адекватного представления типологии объектов необходимо объединять иерархический и прототипический подходы к описанию свойств.

Кроме моделей представления знаний, существуют модели, направленные непосредственно на описание смысла ЕЯ-форм.

Грамматика Монтегю [30] одна из первых логических моделей, направленных на интерпретацию непосредственно естественно-языковых форм. Основным недостатком, не позволяющим добиться адекватного уровня интерпретации, является упрощенный подход к феноменологии естественного языка.

Падежная грамматика Филлмора [24], представляет собой хотя и положительный, но все-таки половинчатый шаг к полной интерпретации семантики предикативных форм. Причиной является то, что инвентарь семантический падежей не охватывает всех типов

семантических ролей актантов из-за несоответствия семантических примитивов в структуре толкования и в когнитивной модели картины мира (см. например [5]).

Теория представления дискурса (Discourse Representation Theory, DRT) [25] - это попытка соединения формальных подходов к описанию грамматик на основе теории Хомского с описанием семантики на основе логики предикатов. В грамматической части DRT проблемой является то, что формальные модели не содержат механизмов для перехода от поверхностного уровня к семантике. В семантической части DRT отсутствует привязка к семантическим типам, существующая, например, в семантических сетях, фреймах, падежной грамматике.

В ряду моделей смысла, а точнее языка в целом, необходимо отметить современные лингвистические теории (модель Смысл<->Текст [10], лексическая семантика [1]), направленные на построение семантической модели естественного языка и языков семантического описания [2,3,21]. Охватывая в целом широкий класс феноменологических явлений языка, эти модели не ориентированы на логику, как средство представления смысла.<sup>2</sup> Картина мира строится средствами естественного языка. Из-за этого возникает необходимость существенной доработки указанных моделей в процессе их дальнейшего инженерного освоения.

В работе [13] предложены подходы к построению грамматики языка и модели смысла, которые лишены указанных недостатков. Описанная далее **семиотическая модель понимания** текста является дальнейшим развитием указанных подходов.

На рис.1 показано схематическое изображение семиотической модели понимания текста.

Понимание текста - сложный процесс, связанный с интерпретацией написанного (сказанного) текста с учетом существования целого комплекса семиотических подсистем. Таких подсистем автором выделяется пять:

- 1) непосредственно фрагмент текста;
- 2) знания агента о языке;
- 3) ситуация, в которую погружено повествование;
- 4) картина мира агента, интерпретирующего текст;

---

<sup>2</sup> Потребность в ориентации на логику вызвана тем, что логика обладает универсальной операциональной семантикой, и ее использование для описания семантического уровня извлекает от "необходимости" придумывать что-то новое.

5) цели агента и связанный с этими целями репертуар моделей поведения (или моделей решения задачи, обеспечивающих достижение цели).

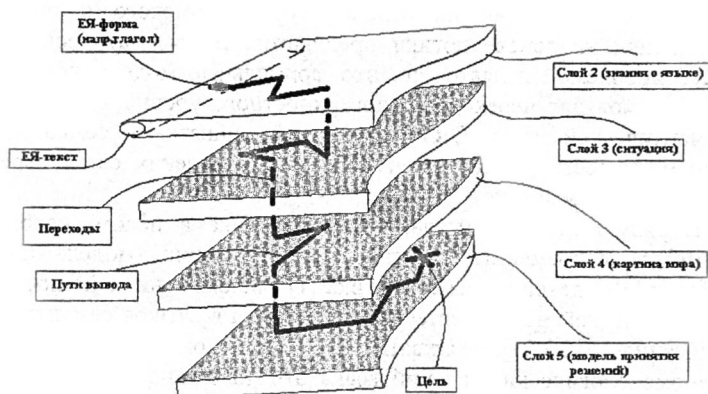


Рис. 1. Семиотическая модель понимания текста

Каждая из этих подсистем является чрезвычайно сложным семиотическим образованием, имеет относительную самостоятельность, подчиняется своим собственным законам функционирования, и, в то же время, является частью общей семиотической системы. Таким образом, понимание рассматривается как совокупность процессов интерпретации смысла фрагмента текста (фразы, предложения, высказывания, абзаца и т.д.) в контексте четырех семиотических подсистем (2-5).

Рассмотрим семантическую природу этих семиотических подсистем подробнее. Существуют целые направления в ИИ [19], которые посвящены моделированию процессов в отдельных подсистемах 1)-5). Так, например, в разделе ИИ "Представление знаний", основной акцент исследований делается на моделирование процессов в подсистемах 3) и 4), т.е. описывается ситуация и картина мира. В разделе ИИ "Экспертные системы" исследуются закономерности принятия решений, т.е. изучается подсистема 5) *Модели принятия решений*. Задачей языкознания является реконструкция представления знаний агента о языке в объеме и качестве, необходимом для успешного функционирования всей семиотической модели понимания текста.

Для более наглядного восприятия этой довольно сложной конструкции прибегнем к схеме, которая, возможно упрощенно, пояснит нам природу взаимосвязей указанных подсистем (Рис.1).

На схеме семиотические подсистема 2)-5) изображены в виде параллельно расположенных слоев, которые на самом деле являются графовыми структурами. Для сохранения единства содержания текста статьи и рисунка нумерация слоев совпадает с нумерацией семиотических подсистем, приведенных ранее. Семиотическая подсистема 1) *Фрагмент текста* схематично изображена в виде стержня, пронизывающего слой 2) *Знания о языке*. Это сделано умышленно, чтобы подчеркнуть тесную взаимосвязь семиотических подсистем 1) и 2), и в то же время - опосредованность связей между подсистемой 1) *Фрагмент текста* и подсистемами 3)-5).

Последнее обстоятельство чрезвычайно важно для уяснения роли языковой модели агента в процессе понимания и принятия решения.

Предложенная схема имеет целый ряд упрощений, которые сделаны для облегчения ее восприятия. Перечислим самые очевидные из этих упрощений.

- Слои представляют собой на самом деле чрезвычайно насыщенные связями графовые структуры (например, для слоя 3) - это семантические сети и графы, типологии объектов и отношений, для слоя 5) - деревья принятия решений и т.д.).

- Связи между слоями не показаны, а их огромное количество, так как именно эти связи являются потенциальными путями переходов.

- В реальной ситуации переходов между слоями гораздо больше и они могут осуществляться в обоих направлениях.

- Связи могут осуществляться не только между соседними слоями, но и в любой последовательности.

- Пути вывода представляют собой весьма разветвленные графовые структуры.

- На схеме не показано то обстоятельство, что при восприятии фрагмента текста ситуацией является контекстное окружение, или, проще говоря, ранее прочитанный текст.

Механизм вывода в семиотической модели понимания текста представляется как процесс поиска путей в рамках одного слоя и переходов между слоями для достижения цели в слое 5). Очевидно, что механизм вывода в такой чрезвычайно сложной семиотической системе требует детальной проработки.

Однако заранее можно сказать, что этот механизм носит смешанный ассоциативно-логический характер, т.е. ряд путей реализуется по ассоциативным связям (как, например, в рамках

модели языка или в ситуационных моделях), другая часть - по логически обусловленным связям (как, например, в рамках модели картины мира или в процессе принятия решений). Также можно сказать, что механизм вывода носит комбинированный последовательно-параллельный характер, так как параллельность вывода осуществляется при поиске путей внутри слоев, а последовательность - за счет переходов между слоями. Единая инвентарная система строительного материала во всех слоях (в первую очередь объектов и отношений между ними) - необходимое условие для осуществления вышеуказанных переходов. Управление выводом (по всем слоям) осуществляется от слоя 5) Модели принятия решений, так как именно этот слой определяет критерии достижения цели.

Отметим, что традиционное разделение частей знаковых систем на синтаксис, семантику и прагматику не исчерпывает вышеприведенной схемы. К примеру, слой 3) *Ситуация* и 5) *Модель принятия решений* можно отнести к прагматике, слой 4) *Картина мира* - к семантике, в слое 2) *Знания о языке* - синтаксис, семантика и прагматика объединены тесной системой взаимосвязей.

Подчеркнем также то, что в семиотической модели понимания текста понимание всегда осуществляется в контексте целеполагания. Это хорошо согласуется с точкой зрения, высказанной в [20].

Последовательность расположения слоев на схеме *Язык-Ситуация-Картина мира-Принятие решения* отражает гипотетическое предположение о типичной последовательности вывода при первичной обработке ЕЯ-феноменов. В качестве альтернативной гипотезы можно рассматривать существование различных комбинаций слоев (2-3-4-5, 2-4-3-5, 2-3-5, 2-4-5, 2-5) в зависимости от ментальной организации агента. Обе эти гипотезы нуждаются в экспериментальной проверке. Итак, знания о языке являются неотъемлемой составной частью семиотической модели понимания текста.

### 3.ПРИБРЕТЕНИЕ ЗНАНИЙ

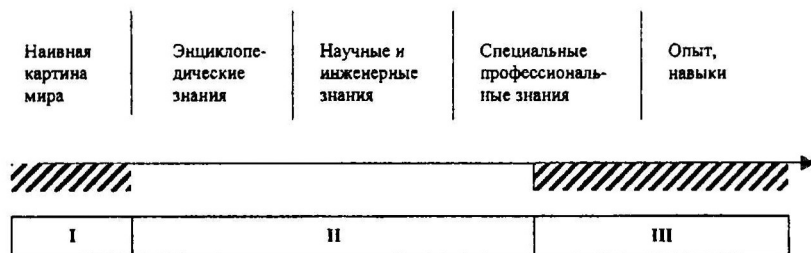
Одновременно с проблемой представления знаний возникла проблема их приобретения. История инженерии знаний в области экспертных систем показала, что "ручная" формализация качественных знаний является чрезвычайно трудоемким процессом [8,12,16]. Как свидетельствуют источники [28,29,31], при построении картины мира в той или иной форме формализовать знания еще труднее.

Существует несколько подходов к проблеме приобретения знаний. Наиболее известны следующие:

- когнитивное моделирование как метод формализации наивной картины мира;
- построение онтологий;
- описание CASE-структур;
- формализация экспертных знаний;
- автоматическое извлечение знаний из текста;
- самообучение в процессе решения задач.

Каждый из перечисленных подходов имеет свою "экологическую" нишу, определяемую видом знаний и участием человека (инженера по знаниям) в процессе приобретения знаний.

Рассмотрим особенности приобретения различных видов знаний на жизненном цикле человека (Рис. 2).

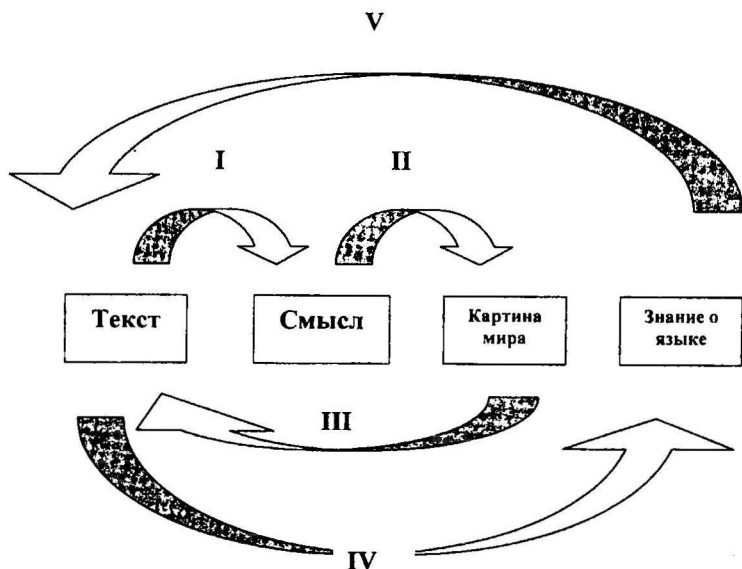


**Рис. 2. Приобретение различных видов знаний на жизненном цикле человека**

Штриховкой на рисунке показаны области, где знания не вербализованы, а, следовательно, процесс их формализации представляет значительные трудности. Заштрихованная область I соответствует наивной картине мира, которая формируется невербальным или полувербальным путем с раннего детства. Эти знания находят свое косвенное отражение в семантике естественного языка. Таким образом, формализацией этой части знаний занимаются разделы ИИ, связанные с построением когнитивных моделей (наивной логики, логики здравого смысла и т.д.) и разделы языкознания, занимающиеся семантикой (лексическая и грамматическая семантика, когнитивная лингвистика в целом). Область II соответствует энциклопедическим знаниям (условно их можно считать школьными), научным и инженерным знаниям (также условно их можно считать

знаниями, полученными в ВУЗе), и части профессиональных знаний. Эти знания сравнительно хорошо вербализованы и описаны в соответствующей учебной литературе, поэтому эта область не заштрихована<sup>3</sup>. Значительная часть профессиональных знаний, а также опыт и навыки, полученные в процессе профессиональной деятельности, в процессе решения задач, не всегда вербализуются. Эти знания получили в литературе название "экспертные знания" и расположены в заштрихованной области III.

Главной проблемой, препятствующей использованию текста для автоматического приобретения знаний, является отсутствие моделей языка и смысла, обеспечивающих стопроцентное понимание. При этом образуется "заколдованный круг" - отсутствие модели языка не позволяет построить картину мира, а отсутствие картины мира - модель языка. Выходом из этой ситуации, по мнению автора, является реализация концепции **частичного понимания**.



**Рис. 3. Цикл обучения в концепции частичного понимания текста**

<sup>3</sup> При этом, конечно, существуют фрагменты знаний в области II, которые также плохо формализованы. В первую очередь - это процедурные знания.



Надо сказать, что большинство моделей языка ориентируются на полное понимание или полный грамматический разбор поступающего на вход лингвистического процессора текста. При этом молчаливо предполагается, что неполное понимание (разбор) приводит к аварийному завершению работы программы. Напротив, концепция частичного понимания предполагает наличие непонимания (неполного разбора) как нормы.

Обучающий цикл "Текст -> Смысл -> Картина\_мира -> Текст -> Знание\_о\_языке -> Текст" (рис. 3) призван реализовать схему обучения по корпусу текстов в рамках концепции частичного понимания. При этом предполагается, что обучение носит итеративный характер. Каждый проход по корпусу текстов добавляет новое знание в картину мира (шаги I, II, III), что в свою очередь позволяет выдвинуть новые гипотезы о семантическом значении грамматических конструкций (шаги IV и V), что в свою очередь позволяет приобретать новое знание и т.д. Этот процесс продолжается до состояния насыщения, при этом новые попытки чтения корпуса текстов не прибавляют знания.

Перечислим условия запуска механизма обучения в рамках предложенной схемы:

- наличие модели смысла, адекватной картине мира и языку;
- наличие структуры представления знаний, адекватной картине мира;
- наличие ядра картины мира, ситуационных моделей и моделей принятия решений;
- наличие ядра знаний о языке;
- наличие специально подобранного корпуса текстов достаточно большого объема.

Несомненно, что описанная схема носит самый общий характер, предложенный подход требует дальнейшей детализации и тщательной проработки.

#### **4. ИСПОЛЬЗОВАНИЕ ЗНАНИЙ**

С точки зрения их использования знания представляют информационную поддержку некоторой интеллектуальной деятельности. Даже если мы имеем дело с информационно-справочной или советующей системой, где, казалось бы, знание существует само по себе, и деятельность семантика существенно отделена от содержательной семантики, всегда возникает необходимость интеллектуальной поддержки диалога с пользователем, что в общем случае является непростой задачей. С

этих позиций интеллектуальную систему **I** можно рассматривать как **n**-ку (**n=6**)<sup>4</sup>:

**I** = <**БЗППР**, **МВБЗППР**, **БЗОКС**, **МВБЗОКС**, **ОМВ**, **ЦЕЛЬ**>, (1)

где

**БЗППР** - база знаний поддержки принятия решений в заданной предметной области (т.е. сценарии CASE-структуры);

**МВБЗППР** - механизм вывода на знаниях поддержки принятия решений; вместе с **БЗППР** - это то, что принято называть опытом;

**БЗОКС** - база знаний, состоящая из общих когнитивных структур, т.е. это то, что составляет багаж общих знаний специалиста (наивная картина мира, школьные, вузовские, энциклопедические и специальные знания);

**МВБЗОКС** - механизм вывода на общих когнитивных структурах;

**ОМВ** - общий механизм вывода;

**ЦЕЛЬ** - условие достижения цели, задающее прагматические характеристики процесса принятия решения.

Необходимость введения в интеллектуальную систему **I** общего механизма вывода и цели можно обосновать следующим образом. При запуске механизмов вывода **МВБЗППР** и **МВБЗОКС** мы получаем чрезвычайно разветвленную сеть результатов вывода (см. рис. 1). В таких условиях нам необходим некоторый механизм координации, в качестве которого выступают - общий механизм вывода (**ОМВ**), и условие остановки, в качестве которого выступает условие достижения цели (**ЦЕЛЬ**).

Работу общего механизма вывода (**ОМВ**) можно описать алгоритмом, изображенным на рис. 4.

Как можно видеть, такой двухуровневый механизм вывода объединяет элементы систем представления знаний и систем принятия решений (экспертных систем) и удовлетворяет следующим условиям:

- подчиненность деятельности определенной цели;
- использование опыта;
- использование знаний.

Традиционные же системы представления знаний не ориентированы на целевое использование знаний и использование опыта, а экспертные системы имеют недостатком отсутствие знаний о картине мира.

---

<sup>4</sup> В систему не включены элементы, связанные с анализом - синтезом естественного языка. Описание ситуаций включено в **БЗППР**.

Старт  
Запуск МВБЗППР.  
Запуск МВБЗОКС.  
 $i = 1, j = 1$

Метка 1. Шаг  $i$  МВБЗППР.

Проверка ЦЕЛИ.

Если ЦЕЛЬ = ЛОЖЬ

Начало\_блока\_1

Формирование условия перехода к  
следующему шагу МВБЗППР: ПЕРЕХОД.

Метка 2. Шаг  $i$  МВБЗОКС.

Проверка ПЕРЕХОДА.

Если ПЕРЕХОД = ЛОЖЬ

Начало\_блока\_2

$j = j + 1$

Возврат к метке 2.

Конец\_блока\_2

Если ПЕРЕХОД = ИСТИНА

Начало\_блока\_3

$i = i + 1$

Возврат к метке 1.

Конец\_блока\_3

Конец\_блока\_1

Если ЦЕЛЬ = ИСТИНА

Выход.

Рис. 4. Алгоритм, описывающий работу  
общего механизма вывода (ОМВ)

## 5. ЗАКЛЮЧЕНИЕ

В настоящей работе исследованы пути дальнейшего развития исследований в области представления, приобретения и использования знаний. В качестве отправной точки принята проблема обработки естественного языка. Рассмотрены достоинства и недостатки моделей представления знаний и смысла ЕЯ-форм, занимающих лидирующие позиции в современных интеллектуальных информационных системах. Описанная далее семиотическая модель понимания текста является дальнейшим развитием указанных подходов. Понимание текста рассматривается как сложный процесс,

связанный с интерпретацией написанного (сказанного) текста с учетом существования целого комплекса семиотических подсистем. К таким подсистемам относятся: непосредственно фрагмент текста; знания агента о языке; ситуация, в которую погружено повествование; картина мира агента, интерпретирующего текст; цели агента и связанный с этими целями репертуар моделей поведения (или моделей решения задачи, обеспечивающих достижение цели). В работе предложена концепция частичного понимания текста. Для приобретения знаний в рамках этой концепции описан цикл обучения, позволяющий объединить процесс автоматического формирования картины мира и знаний о языке по корпусу текстов. Для целей использования знаний предложен двухуровневый механизм вывода, который объединяет элементы систем представления знаний и систем принятия решений (экспертных систем) и удовлетворяет следующим условиям: подчиненность деятельности определенной цели; использование опыта; использование знаний. Работу можно рассматривать как программу научных исследований, которую, по мнению автора, предстоит реализовать научному сообществу в ближайшее десятилетие.

## ЛИТЕРАТУРА

[1] Апресян Ю.Д. *Лексическая семантика. Синонимические средства языка*. М., Наука, 1974, 367 с.

[2] Апресян Ю.Д. Избранные труды, том II. *Интегральное описание языка и системная лексикография*. - М.: Школа "Языки русской культуры", 1995. -767 с.

[3] Вежбицкая Анна. *Семантические универсалии в описании языков*. /Пер. с англ. А.Д. Шмелева под ред. Т.В. Булыгиной/ М.: Школа "Языки русской культуры", 1999. XII, -780 с.

[4] Дьячко А.Г., Конвисар Е.П., Поляков В.Н. и др. *Извлечение и сбор научных данных в компьютерной сети Интернет*. // КИИ'98. Труды конференции. - Пушино, 1998. С. 466-473.

[5] Кандрашина Е.Ю., Литвинцева Л.В., Поспелов Д.А. *Представление знаний о времени и пространстве в интеллектуальных системах*. - М.: Наука, 1989, -328 с.

[6] Кобозева И.М. *Как мы описываем пространство, которое видим: типы и ранги объектов*. Труды Международного семинара "Диалог'96" по компьютерной лингвистике и ее приложениям. Пушино, 4-9 мая 1996, -М: 1996, с. 109-111.

[7] Кривоносов А.Т. *К взаимоотношению языка и мышления. (логическое умозаключение в естественном языке)* /Обработка текста

и когнитивные технологии: Сборник /Под ред. Соловьева В.Д., - Пушкино:1999, (Вып 3)

[8]Ларичев О.И., Мошкович Е.М. *Качественные методы принятия решений. Вербальный анализ решений.* - М.: Наука. Физматлит, 1996.

[9]Минский М. Фреймы и представление знаний. - М.: Энергия, 1979, -150 с.

[10]Мельчук И.А. *Опыт теории лингвистических моделей "Смысл <-> Текст".* -М.: Наука, 1974.

[11]*Логические проблемы знания.* Автор неизвестен (на электронных носителях).

[12]Осипов Г.С. *Приобретение знаний интеллектуальными системами.* М.: Наука, Физматлит., 1997, 112 с.

[13]Поляков В.Н. *Модели алгоритмического типа для распознавания семантических связей в системах машинной обработки естественного языка.* Автореферат дисс. на соиск. уч. ст. к.т.н., М., Изд-во МИСиС, 1997, 20 с.

[14]Поляков В.Н. *Перспективы интеллектуального освоения информационного пространства сети ИНТЕРНЕТ научным сообществом /Труды международного семинара Диалог'99 по компьютерной лингвистике и ее приложениям.* Таруса, 1999 г., т.2, с.237-247.

[15]Поляков В.Н. *К когнитивной модели русского глагола.* // Обработка текста и когнитивные технологии. Сборник / Под ред. Соловьева В.Д., - Пушкино:, 1999, Вып. 4 (В печати).

[16]Попов Э.В. *Экспертные системы. Решение неформализованных задач в диалоге с ЭВМ.*-М.:Наука,1987.-288 с.

[17]Поспелов Д.А. *Данные и знания.* // Справочник. Искусственный интеллект., кн.2., Модели и методы., - М., Радио и связь., 1990, с.7-14.

[18]Ришар Ж.Ф. *Ментальная активность. Понимание, рассуждение, нахождение решений.* Сокр. пер. с франц. Т.А. Ребеко М. Издательство "Институт психологии РАН". 1998 г.-232 с.

[19]Справочник. *Искусственный интеллект.*, в 3-х кн., - М., Радио и связь., 1990.

[20]Сулейманов Д. Ш. *Обработка ЕЯ-текстов на основе прагматически-ориентированных лингвистических моделей.* // Обработка текста и когнитивные технологии: Сборник / Под ред. Соловьева В.Д., - Пушкино:, 1999, (Вып 3), с 205-212.

[21]Шалапина З.М. *Об одном формализме для записи толкований слов и словосочетаний.* // Обработка текста и когнитивные технологии: Сборник /Под ред. Дьячко А.Г., - Москва, Пушкино: ОНТИ ПНЦ РАН, 1997, (Вып.1), с 73-110.

[22]Berkeley E.C. *Symbolic Logic and Intelligent Machines*. Reinold Publishing Corporation, New York. Перевод: Э.Беркли. *Символическая логика и разумные машины*. М.: ИЛ, 1961, 260.

[23]Conceptual Graphs Home Page <http://concept.cs.uah.edu/CG/>

[24]Fillmore Ch.T. *The case for case.*// Universal in linguistic theory. New York, 1968. (Перевод: "Новое в зарубежной лингвистике, вып.10, М., Прорпеец, 1981,с.369-495)

[25]Kamp H. and Reyle U. *From Discourse to Logic*. Kluwer Academic Press, 1993.

[26]Knowledge                      Interchange                      Format                      (KIF)  
<http://logic.stanford.edu/kif/kif.html>

[27]Lehmann F.W. *Semantic Networks* // Computers & Mathematics with Applications, V.23,N.2-5,1992.

[28]Lehnert Wendy G., Dyer Michael G., Johnson Peter N. Yang C.J. and Harley Steve. BORIS - *An Experiment in-Depth Understanding of Narratives*. - "Artificial Intelligence", vol.20, N 1, 1983, p.15-62

[29]Melchuk I.A. *The Russian Language in the Meaning - Text Perspective*. Wiener Slawistischer Almanach, Sonderband, 39, Moskau-Wien,1995.

[30]Montague Richard, *Universal Grammar, Theoria*, 36,373-398

[31]Schank R.. *Conceptual Information Processing*. New York, American Elsevier., 1975.Перевод: Шэнк Р. *Обработка концептуальной информации.*: Пер. с англ.-М.:Энергия,1980.

[32]Survey of the State of the Art in Human Language Technology / Cole, Ronald, et al (eds.) *Studies in Natural Language Processing* . Cambridge University Press 1998. 533 pp.

[33]Thayse A. et all. *Approche logique de l'intelligence artificielle*. 1.De la logique classique a la programmation logique., Bordas, Paris, 1988. Перевод: *Логический подход к искусственному интеллекту*. 1. От классической логики к логическому программированию. Пер. с фр. / Тейз А.,Грибомон П.,Луи Ш. и др., - М.:Мир,1990,-132 с.